



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------|-------------|----------------------|---------------------|------------------|
| 09/866,394 | 05/25/2001 | Nevenka Dimitrova | US 010265 | 5012 |

24737 7590 10/21/2005

PHILIPS INTELLECTUAL PROPERTY & STANDARDS
P.O. BOX 3001
BRIARCLIFF MANOR, NY 10510

| |
|----------|
| EXAMINER |
|----------|

ZHOU, TING

| | |
|----------|--------------|
| ART UNIT | PAPER NUMBER |
|----------|--------------|

2173

DATE MAILED: 10/21/2005

Please find below and/or attached an Office communication concerning this application or proceeding.



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

MAILED

OCT 21 2005

Technology Center 2100

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Application Number: 09/866,394
Filing Date: May 25, 2001
Appellant(s): DIMITROVA ET AL.

Michael A. Scaturro
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 28 July 2005 appealing from the Office action mailed

25 January 2005.

JD

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

- Dimitrova et al., "Color SuperHistogram for Video Representation" IEEE (1999).
- 5,805,733 Dimitrova et al. 9-1998

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claims 1-5, 7-15, 17-24, 26-33 and 35-38 are rejected under 35 U.S.C. 103(a) as being unpatentable over the article entitled "Color SuperHistograms for Video Representation", written by Dimitrova et al., and Wang et al. U.S. Patent 5,805,733.

Referring to claims 1, 11, 21 and 30, Dimitrova et al. teach an apparatus, system, method and computer executable instructions comprising a visual summary controller capable of creating a visual summary of video material (Dimitrova et al.: page 316, Figure 1), wherein the visual summary controller is capable of extracting frame signatures (histograms) from keyframes of video material and capable of using the frame signatures to create superhistograms from the keyframes (Dimitrova et al.: page 314, right column, lines 11-25, page 315, section 2 and page 316, section 2.3; this is further shown in Figure 1). However, although Dimitrova et al. teach using the frame signatures and superhistograms to create a visual summary of video material in a broad sense (representing video segments by computing superhistograms) (Dimitrova et al.:

Art Unit: 2173

Abstract), Dimitrova et al. fail to explicitly teach selecting representative keyframe images for each superhistogram to create a compact visual summary of the video material, wherein the representative images include at least one of the first image in each family histogram, the most meaningful image in each superhistogram, a randomly chosen image and an image that is closest to the cluster center. Wang et al. teach the analysis of scenes and frames in video materials (Wang et al.: column 1, lines 53-56 and Figure 2) similar to that of Dimitrova et al. In addition, Wang et al. further teach selecting representative keyframe images from each group of related scenes to create a compact visual summary of the video material (summarizing a video sequence by taking one representative frame from each set of related scenes with similar average color histograms, to represent the set to enable the user to view a large sampling of video sequence images) (Wang et al.: column 1, lines 51-67 and column 2, lines 1-24; this is further shown in Figure 3), wherein the representative images include at least one of the first image in each family histogram, the most meaningful image in each superhistogram, a randomly chosen image and an image that is closest to the cluster center (the representative frame image can be taken from the temporally medial scene in the set or from one of the frames of the longest scene in the set of related scenes) (Wang et al.: column 3, lines 37-66). It would have been obvious to one of ordinary skill in the art, having the teachings of Dimitrova et al. and Wang et al. before him at the time the invention was made, to modify the visual summary controller capable of extracting frame signatures from keyframes to create superhistograms of Dimitrova et al., to include the further step of selecting representative keyframes from those superhistograms and using the representative keyframe images to create a compact visual summary, taught by Wang et al. One would have been motivated to make such a combination in order to meet the need of being able

Art Unit: 2173

to readily access and manipulate video information, by cataloguing and storing the potentially thousands of hours of video for rapid future retrieval, browsing and use, created by the increasing availability and use of digital video and the increasing integration of computer technologies and video production technologies.

Referring to claims 2, 12, 22 and 31, Dimitrova et al. teach the filtering of keyframes (merging of histograms into family histograms) and extracting frames signatures (computing color histograms) from the filtered keyframes before using the frame signatures (histograms) to create the superhistogram representing a visual summary of the video material (page 315, right column, section 2 and page 316, left column, section 2.3).

Referring to claims 3, 13, 23 and 32, Dimitrova et al. teach the use of superhistograms to cluster the filtered keyframes (the ordered merging of the family histograms to create the superhistogram), wherein the clustered keyframes (superhistogram) represents the visual summary of the video material, as recited on page 314, right column, lines 11-25 and shown in Figure 1.

Referring to claims 4 and 14, Dimitrova et al. teach the use of a histogram as the frame signature used to compute superhistograms (page 314, right column, lines 11-15).

Referring to claims 5, 15, 24 and 33, Dimitrova et al. the use of the L1 distance measure method, L2 distance measure method, histogram intersection method, Chi-Square test and Bin-wise histogram intersection method to computer the histogram difference (page 315, right column).

Referring to claims 7, 17, 26 and 35, Dimitrova et al. teach the ability to select the family histograms (the top n largest families) to use to create the superhistogram used to create the visual summary (page 316, section 2.4).

Referring to claims 8, 18, 27 and 36, while Dimitrova et al. teach all of the limitations as applied to claims 1, 11, 21 and 30 above, Dimitrova et al. fail to explicitly teach the capability to retrieve a visual summary stored in a memory unit and causing the visual summary to be displayed in response to a user request. Wang et al. teach the analysis of scenes and frames in video materials (Wang et al.: column 1, lines 53-56 and Figure 2) similar to that of Dimitrova et al. In addition, Wang et al. further teach the capability of letting a user select a visual summary for viewing, retrieving that visual summary from memory and displaying it in response to the user's request (displaying visual summaries of scenes in a movie bar and allowing users to access the summaries by selecting the segments corresponding to the scenes) (Wang et al.: column 2, lines 16-29 and shown in Figures 2 and 3). It would have been obvious to one of ordinary skill in the art, having the teachings of Dimitrova et al. and Wang et al. before him at the time the invention was made, to modify the visual summary controller capable of extracting frame signatures from keyframes to create superhistograms of Dimitrova et al., to include the retrieval and display of the visual summary in response to a user request, as taught by Wang et al. One would have been motivated to make such a combination to give users the flexibility to select which scenes to watch, saving them time from having to browse through all of the other irrelevant scenes; furthermore, because the increasing availability and use of digital video and the increasing integration of computer technologies and video production technologies have produced the need to be able to readily access and manipulate video information, it would have

been advantageous to make such a combination in order to provide users a way to summarize the content of video quickly and easily, in order to catalogue and store the potentially thousands of hours of video for rapid future retrieval, browsing and use.

Referring to claims 9, 19, 28 and 37, Dimitrova et al. teach the use of the visual summary obtained from the superhistograms to access at least a portion of the video material (classifying and searching in video archives), as recited on page 317, section 4.2.

Referring to claim 10, 20, 29 and 38, while Dimitrova et al. teach all of the limitations as applied to claims 1, 11, 21 and 30 above, Dimitrova et al. fail to explicitly teach the creation of new video material using the compact visual summaries. Wang et al. teach the analysis of scenes and frames in video materials (Wang et al.: column 1, lines 53-56 and Figure 2) similar to that of Dimitrova et al. In addition, Wang et al. further teach the creation of new video material using the compact visual summaries (a collage made up of representative frames for each set of summarized scenes) (Wang et al.: column 3, lines 53-57). It would have been obvious to one of ordinary skill in the art, having the teachings of Dimitrova et al. and Wang et al. before him at the time the invention was made, to modify the visual summary controller capable of extracting frame signatures from keyframes to create superhistograms of Dimitrova et al., to include the creation of new video material, as taught by Wang et al. It would have been advantageous for one to utilize such a combination in order to conserve processor time and storage space by utilizing the already existing visual summaries in the creation of new visual materials; furthermore, because the increasing availability and use of digital video and the increasing integration of computer technologies and video production technologies have produced the need to be able to readily access and manipulate video information, it would have been advantageous

to make such a combination in order to provide users a way to summarize the content of video quickly and easily, in order to catalogue and store the potentially thousands of hours of video for rapid future retrieval, browsing and use.

(10) Response to Argument

The applicant argues that Wang's method for selecting representative frames is based on a temporal ordering of frames, whereas in contrast, the method of the invention selects representative frames based on a non-temporal order, through the use of certain terms throughout the claims and defined in the specification; specifically, the terms that are indicative of a non-temporal ordering include: "superhistogram", "family histogram", and "cluster center". The examiner respectfully disagrees with the applicant's assertion that terms used in the claim language, i.e. "superhistogram", "family histogram" and "cluster center" are defined to be non-temporal based in the specification. The examiner respectfully argues that there is no basis or support in the specification or claim language for the applicant's argument. Furthermore, in response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., non-temporal based ordering) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). However, in the interest of furthering prosecution, the examiner has carefully read the specification and found no definition of terms such as "superhistogram", "family histogram" and "cluster center" to be non-temporal; in fact, there were no mentioning of non-temporal ordering, let alone exclusion of selecting

Art Unit: 2173

representative frames based on a temporal ordering or specification of selecting representative frames based on a non-temporal ordering. The examiner only found one passage in the specification (pages 12, line 19 - page13, line 8) referring to terms such as “superhistogram” and “family histogram” with respect to time:

“Superhistogram application 240 computes superhistograms by computing color histograms for individual shots and then merging the histograms into a single cumulative histogram called a family histogram based on a comparison measure. A family histogram originally represents the color union of two shots. As new frames are added, the family histogram accumulates the new colors from the respective shots. If a histogram of a new frame differs from the family histograms previously constructed, then a new family histogram is formed. An entire television program, for example, may be represented by a few family histograms. **The set of family histograms is ordered with respect to the length of the temporal segment of video that they represent.** The ordered set of family histograms is called a superhistogram.”

As shown by the specification passage above, not only is there no exclusion of temporal-based ordering, the applicant’s specification actually states that the family histograms are ordered according to a temporal element. Therefore, the examiner respectfully disagrees with the applicant’s assertion that Wang’s method for selecting representative frames based on a temporal ordering of frames is in contrast to the invention’s **defined** and **claimed** method of non-temporal based order.

The applicant argues that Wang does not teach that the representative image includes at least one of the first image in each family histogram, the most meaningful image in each superhistogram, a randomly chosen image and an image that is closest to the cluster center; specifically, the applicant argues that Wang fails to teach the representative image includes the

Art Unit: 2173

most meaningful image in each superhistogram, nor an image that is closest to the cluster center. The examiner respectfully disagrees. First with reference to the representative image being an image that is closest to the cluster center, Wang teaches that the representative frame can be a frame that is halfway between the first and last scene, as recited in column 2, lines 12-15 and column 3, lines 57-59. The applicant argues that a critical distinction between the method of Wang and the method of the invention is that in accordance with the method of the invention, the cluster center is derived as a non-temporal ordering. As the examiner stated above, the applicant fails to provide any basis or support for such as an argument in either the claim language themselves or in the body of the specification. Since the specification of the applicant's invention fails to include a definition of "cluster center", taking the broadest reasonable interpretation, one of ordinary skill in the art would interpret an image that is closest to the cluster center to be an image that is closest to the center, or middle of a cluster of scenes. Since Wang teaches that the representative frame is a frame that is halfway between the first and last scenes, Wang teaches that the representative frame is at the center of the cluster of frames and is therefore a frame that is closest to the cluster center. Second with reference to the representative image being the most meaningful image in each superhistogram, Wang teaches that the representative frame can be a frame taken from the longest scene, since the longest scene is most indicative of the content of the related scenes, as recited in column 3, lines 59-62. The applicant argues that simply selecting one frame from among the frames of the longest scene, as taught by Wang, does not teach or disclose the "most meaningful frame in the group" since the specification of the invention recites with particularity (by example) what constitutes the "most meaningful frame", i.e. a person's face, an important text, etc. The examiner respectfully

disagrees. In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., the representative frame being a person's face, an important text, etc.) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). Although the applicant's specification gives **examples** of some frames that can be considered the "most meaningful frame" (the specification recites at page 16, lines 21-22, "The term 'meaningful image' **may** refer to a frame with a person's face, an important text, **etc.**"), it does not require that the most meaningful frame **has to be** a frame with a person's face or an important text. Wang teaches selecting representative scenes for presentation to a user that have the "most significant content", as recited in column 5, lines 11-20. Wang further recites that a representative frame "can be taken as one of the frames of the longest scene in a set, the longest scene being most indicative of the content of the related scenes", in column 3, lines 59-62. Therefore, the examiner respectfully argues that since the longest scene is most indicative of the content of the related scenes, the longest scene is the most meaningful of the group.

In response to applicant's arguments and example that the method according to the invention for calculating a representative keyframe image that is the most meaningful frame in the group or an image that is closest to the cluster center will produce results that are in sharp contrast to those produced according to the method of Wang, the fact that the results of the inventions are different cannot be the basis for patentability; the results of the applicant's

Art Unit: 2173

invention are not claimed, it is the process of producing the results that are claimed and as long as the prior art teaches every step of the process as claimed, the fact that the final results may be different are irrelevant. See *Ex parte Obiaya*, 227 USPQ 58, 60 (Bd. Pat. App. & Inter. 1985).

The language of the claims recite "wherein said representative images include at least one of (1) the first image in each family histogram, (2) the most meaningful image in each superhistogram, (3) a randomly chosen image, and (4) an image that is closest to the cluster center". From the above responses to the applicant's arguments, the examiner respectfully asserts that Wang teaches the representative image includes at least two of the four images claimed in the recited claims, specifically Wang teaches the representative image can be both the most meaningful image in each superhistogram or an image that is closest to the cluster center.

(11) Related Proceeding(s) Appendix

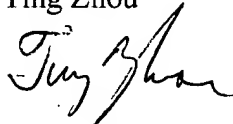
No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,


Art Unit: 2173

Ting Zhou



Conferees:

John Cabeca


JOHN CABECA
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100

Kristine Kincaid


KRISTINE KINCAID
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100